UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/685,656	10/14/2003	W. Todd Daniell	030220; 190250-1300	5661

38823 7590 11/23/2009

AT&T Legal Department - TKHR
Attn: Patent Docketing
One AT&T Way
Room 2A-207
Bedminster, NJ 07921

EXAMINER	
MACILWINEN, JOHN MOORE JAIN	
ART UNIT	PAPER NUMBER
2442	

MAIL DATE	DELIVERY MODE
11/23/2009	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary	Application No.	Applicant(s)	
	10/685,656	DANIELL ET AL.	
	Examiner	Art Unit	
	John M. MacIlwinen	2442	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE *3* MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *10 September 2009*.

2a)☒ This action is **FINAL**. 2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1,6,11-14,16,17 and 19-39* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1,6,11-14,16,17 and 19-39* is/are rejected.

7)☒ Claim(s) *23-25 and 30-32* is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☒ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☐ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date *See Continuation Sheet*.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

Continuation of Attachment(s) 3). Information Disclosure Statement(s) (PTO/SB/08), Paper No(s)/Mail Date :6/25/09,7/28/09,8/14/09,9/14/09,10/22/09, 11/17/2009.

DETAILED ACTION

Response to Arguments

1. Applicant's arguments filed 9/10/2009 have been fully considered

2. Applicant argues the rejections made under 35 USC 112, first paragraph. Applicant argues that regarding the Written Description rejection of claim 6 for the phrase “the displaying characters of the SMTP email address”; specifically, Applicant argues that page 18 recites a “displayable body of characters”. However, a “displayable body of characters” is not the same thing as “the displaying characters of the SMTP email address”. Applicant’s argument thus is not persuasive.

3. Applicant continues arguing that “the displaying characters of the SMTP address” is clear as “this phrase clearly indicates those characters that are being displayed”. However, in the context of Applicant’s claim language and specification, which discusses the general non-displaying characters, it is unclear what Applicant intends to specify through the use of this phrase (that is, are there other, non-displaying characters of the SMTP address? How is this language in particular applicable to an SMTP address? etc.). Applicant’s lack of written description for this phrase exacerbates this issue. Applicant’s arguments are not persuasive.

4. Regarding the rejection of claims 24 and 31 under 35 USC 101, Applicant argues that in said claims "the 'means for' terminology . . . must include hardware.” In light of Applicant’s clarification, Applicant’s arguments are persuasive and the rejection has been dropped.

5. Regarding the arguments addressing the rejections under 35 USC 103, Applicant's argument Applicant's arguments have been fully considered. Said arguments generally fail to comply with 37 CFR 1.111(b) because they amount to a general allegation that the claims define a patentable invention without specifically pointing out how the language of the claims patentably distinguishes them from the references. Applicant's arguments merely recite the amended claim language and state that individual references "fail to even suggest" said claim language.

6. However, in view of the arguments against Gordon and the clarification provided by the amended claim language, a new grounds of rejection and objections have been made which is discussed further below.

Claim Objections

7. Claim 23 is objected to because of the following informalities: said claim recites "the non-displaying control characters" on line 9; there is insufficient antecedent basis for this limitation.

8. Claim 24 is objected to because of the following informalities: said claim recites "the non-displaying comments and the non-displaying control characters" on lines 6 - 7; there is insufficient antecedent basis for this limitation.

9. Claim 25 is objected to because of the following informalities: said claim recites "the non-displaying comments and the non-displaying control characters" on lines 7 - 8; there is insufficient antecedent basis for this limitation.

10. Claim 30 is objected to because of the following informalities: said claim recites "the non-displaying comments and the non-displaying control characters" on lines 7 - 8; there is insufficient antecedent basis for this limitation.

11. Claim 31 is objected to because of the following informalities: said claim recites "the non-displaying comments and the non-displaying control characters" on lines 5 - 6; there is insufficient antecedent basis for this limitation.

12. Claim 32 is objected to because of the following informalities: said claim recites "the non-displaying comments and the non-displaying control characters" on lines 6 - 7; there is insufficient antecedent basis for this limitation.

Appropriate correction is required.

Specification

13. The specification is objected to as failing to provide proper antecedent basis for the claimed subject matter for the reasons given below in the 35 USC 112 written description rejection. See 37 CFR 1.75(d)(1) and MPEP § 608.01(o).

Claim Rejections - 35 USC § 112

14. The following is a quotation of the first paragraph of 35 U.S.C. 112:

> The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same and shall set forth the best mode contemplated by the inventor of carrying out his invention.

15. Claim 6 recites "the displaying characters of the STMP email address"; there is a lack written description for said limitation.

16. The following is a quotation of the second paragraph of 35 U.S.C. 112:

> The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

17. Claims 1, 6 and 39 are rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention.

18. Regarding claim 1, said claim recites, beginning on line 19,

"(D3) if the character is a space, determine whether the space is adjacent to a solitary 'i' or 'a'; and

(D4) if the non-alphabetic character is not a space, filtering the determined non-alphabetic displaying characters from the email;"

It is unclear what the relationship between the above steps is, and what step the purpose is of D3 as the determination of step D3 is never utilized. Step D3 is generally indefinite and unclear and makes the entirety of step "(D)" of claim 1, beginning on line 14, indistinct as it is unclear and indistinct what the relationship is between the sub-steps (D1) through (D4).

19. Regarding claim 39, said claim is similarly indistinct and unclear as it relies on the indistinct and unclear language of claim 1 for meaning and scope.

20. Regarding claim 6, said claim recites "a token representative of the displaying characters of the SMTP email address". It is unclear what "the displaying characters of the STMP email address" refers to.

Claim Rejections - 35 USC § 103

21. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

22. Claims 6, 11-14, 16, 17 and 19 – 38 are rejected under 35 U.S.C. 103(a) as being unpatentable over Shipp (US 2004/0093384 A1) in view of Milliken (US 2004/0073617 A1), Sahami (A Bayesian Approach to Filtering Junk E-Mail) and Woitaszek (Identifying Junk Electronic Mail in Microsoft Outlook with a Support Vector Machine).

23. Regarding claim 6, Shipp shows a method comprising receiving, at a computing device, an email message comprising a text body ([64,65]), an SMTP email address ([39,43,69]), and a domain name corresponding to the SMTP email address ([39,45,46]), the text body including displaying characters ([64-67]) and non-displaying characters ([57-58, 61-73]);

tokenizing the SMTP email address to generate a token representative of the SMTP email address ([39,43,63])

tokenizing the domain name to generate a token representative of the domain name ([22]), and determining a spam probability value from the generated tokens ([14,76]).

Shipp does not show tokenizing the attachment to generate a token that is representative of the attachment nor does Shipp show searching for the non-displaying

characters in the email and removing the searched non-displaying characters, including non-displaying comments and non-displaying control characters.

Milliken shows tokenizing the attachment to generate a token that is representative of the attachment ([10-13 and 51 – 53]) and

searching for the non-displaying characters in the email and removing the searched non-displaying characters, including non-displaying comments and non-displaying control characters ([69]).

It would have been obvious to one of ordinary skill in the art at the time of the invention to modify the disclosure of Shipp with that of Milliken in order to better identify spam email, as at the time of Shipp's disclosure, spam email was thought “currently” not to be associated with attachments ([81]); spam and attachments are however an area for which Milliken’s more recent disclosure provides updated guidance.

Shipp in view of Milliken do not show explicitly show where the tokens are sorted in accordance with the corresponding determined spam probability value to determine a predefined number of interesting tokens, the predefined number of interesting tokens being a subset of the generated tokens.

Sahami shows selecting a predefined number of interesting tokens, the predefined number of interesting tokens being a subset of the generated tokens (pg. 4, col. 1, showing having initially “several thousand” features, then selecting 500 of said features after first sorting out features that occur fewer than 3 times (pg. 4, col. 2) and then selecting, of the remaining feature, the 500 features with the highest non-neutral probability value (pg. 6, col. 1, paragraph 1)).

It would have been obvious to one of ordinary skill in the art at the time of the invention to modify the disclosure of Shipp in view of Milliken with that of Sahami in order to more accurately identify spam email (Sahami, Abstract).

Shipp in view Milliken and Sahami thus do show selecting a subset of the generated tokens based on probability value as well as where the interesting tokens are a subset of the generated tokens (Sahami, pg. 6, col. 1, paragraph 1), but do not show explicitly show where the tokens are sorted in accordance with the corresponding determined spam probability value.

Woitaszek shows where the tokens are sorted in accordance with the corresponding determined spam probability value (Tables 4 and 5).

It would have been obvious to one of ordinary skill in the art at the time of the invention to modify the disclosure of Shipp in view of Milliken and Sahami with that of Woitaszek in order to arrange the calculated values in a logical manner, enabling a simple method of extracting the most interesting results (as discussed by Sahami) via simply taking the top occurring results in Woitaszek's sorted list, as well as to include the abilities to integrate the spam software into a commonly used email program (Woitaszek, Abstract, pg. 1 col. 2).

24. Regarding claim 11, Shipp in view of Milliken, Sahami and Woitaszek further show assigning a spam probability value to the token representative of the SMTP email address (Shipp [18,23,39,40-43], Woitaszek, Tables 4 and 5) and

assigning a spam probability value to the token representative of the domain name (Shipp [22]).

and generating a Bayesian probability values using the spam probability values assigned to the tokens (Sahami, pg.2, col. 2; pg. 4, col. 2; pg. 6, col. 1).

25. Regarding claim 12, Shipp in view of Milliken, Sahami and Woitaszek further show comparing the generated Bayesian probability value with a predefined threshold value (Sahami, pg.2, col. 2; pg. 4, col. 2; pg. 6, col. 1).

26. Regarding claim 13, Shipp in view of Milliken, Sahami and Woitaszek further show categorizing the email message as spam in response to the Bayesian probability value being greater than the predefined threshold (Sahami, pg.2, col. 2; pg. 4, col. 2; pg. 6, col. 1).

27. Regarding claim 14, Shipp in view of Milliken, Sahami and Woitaszek further show categorizing the email message as non-spam in response to the Bayesian probability value being not greater than the predefined threshold (Sahami pg. 6 col. 1).

28. Regarding claim 16, Shipp in view of Milliken, Sahami and Woitaszek further show receiving an email message including a text body (Shipp [64,65]).

29. Regarding claim 17, Shipp in view of Milliken, Sahami and Woitaszek further show tokenizing the words in the text body to generate tokens representative of the words in the text body (Shipp [64,65]).

30. Regarding claim 19, Shipp in view of Milliken, Sahami and Woitaszek further show assigning a spam probability value to each of the tokens representation of the words in the text body (Woitaszek, Tables 4 and 5)

assigning a spam probability value to token representative of the attachment (Woitaszek, Tables 4 and 5, and Milliken, [10-13]),

and generating a Bayesian probability value using the spam probability values assigned to the token (Sahami, pg. 4 col. 2).

31. Regarding claim 20, Shipp in view of Milliken, Sahami and Woitaszek further show comparing the generated Bayesian probability value with a predefined threshold value (Sahami, pg. 4 col. 2).

32. Regarding claim 21, Shipp in view of Milliken, Sahami and Woitaszek further show categorizing the email message as spam in response to the Bayesian probability value being greater than the predefined threshold (Sahami, pg.2, col. 2; pg. 4, col. 2; pg. 6, col. 1).

33. Regarding claim 22, Shipp in view of Milliken, Sahami and Woitaszek further show categorizing the email message as non-spam in response to the Bayesian probability value being not greater than the predefined threshold (Sahami, pg. 6 col. 1).

34. Regarding claim 23, Shipp shows a memory component that stores at least the following:

email receive log configured to receive an email message comprising a text body (Shipp, [64,65]), an SMTP email address (Shipp, [39-43,69]), and a domain name corresponding to the SMTP email address and an attachment, (Shipp, [39,45,46]) the email message further including (Shipp, [64-67]) and non-displaying characters (Shipp, [57-58, 61-73]);

tokenizing logic configured to tokenize the SMTP email address to generate a token representative of the SMTP email address (Shipp, [39,43,63])

tokenizing the domain name to generate a token representative of the domain

name (Shipp, [22]), and

determining a spam probability value from the generated tokens (Shipp, [14,76]).

Shipp does not show tokenizing the attachment to generate a token that is representative of the attachment nor does Shipp show searching logic that is configured to search for the non-displaying characters in the email and removing logic that is configured to remove the searched non-displaying characters, including non-displaying comments and the non-displaying control characters, wherein only the displaying characters are tokenized.

Milliken shows tokenizing the attachment to generate a token that is representative of the attachment ([10-13 and 51 – 53]) and

searching logic that is configured to search for the non-displaying characters in the email and removing logic that is configured to remove the searched non-displaying characters, including non-displaying comments and the non-displaying control characters ([69]), wherein only the displaying characters are tokenized ([69]).

It would have been obvious to one of ordinary skill in the art at the time of the invention to modify the disclosure of Shipp with that of Milliken in order to better identify spam email, as at the time of Shipp's disclosure, spam email was thought “currently” not to be associated with attachments ([81]); spam and attachments are however an area for which Milliken’s more recent disclosure provides updated guidance.

Shipp in view of Milliken do not show explicitly show where the tokens are sorted in accordance with the corresponding determined spam probability value to determine a predefined number of interesting tokens, the predefined number of interesting tokens

being a subset of the generated tokens.

Sahami shows selecting a predefined number of interesting tokens, the predefined number of interesting tokens being a subset of the generated tokens (pg. 4, col. 1, showing having initially "several thousand" features, then selecting 500 of said features after first sorting out features that occur fewer than 3 times (pg. 4, col. 2) and then selecting, of the remaining feature, the 500 features with the highest non-neutral probability value (pg. 6, col. 1, paragraph 1)).

It would have been obvious to one of ordinary skill in the art at the time of the invention to modify the disclosure of Shipp in view of Milliken with that of Sahami in order to more accurately identify spam email (Sahami, Abstract).

Shipp in view Milliken and Sahami thus do show selecting a subset of the generated tokens based on probability value as well as where the interesting tokens are a subset of the generated tokens (Sahami, pg. 6, col. 1, paragraph 1), but do not show explicitly show where the tokens are sorted in accordance with the corresponding determined spam probability value.

Woitaszek shows where the tokens are sorted in accordance with the corresponding determined spam probability value (Tables 4 and 5).

It would have been obvious to one of ordinary skill in the art at the time of the invention to modify the disclosure of Shipp in view of Milliken and Sahami with that of Woitaszek in order to arrange the calculated values in a logical manner, enabling a simple method of extracting the most interesting results (as discussed by Sahami) via simply taking the top occurring results in Woitaszek's sorted list, as well as to include

the abilities to integrate the spam software into a commonly used email program (Woitaszek, Abstract, pg. 1 col. 2).

35. Regarding claim 24, Shipp shows means for receiving an SMTP email address, and a domain name corresponding to the SMTP email address (Shipp, [39,45,46]) and an address (Shipp, [39,45,46]) the email message further including (Shipp, [64-67]) and non-displaying characters (Shipp, [57-58, 61-73]);

means for tokenizing the SMTP email address to generate a token representative of the SMTP email address (Shipp, [39,43,63])

means for tokenizing the domain name to generate a token representative of the domain name (Shipp, [22]), and

means for determining a spam probability value from the generated tokens (Shipp, [14,76]).

Shipp does not show all of means for tokenizing the attachment to generate a token that is representative of the attachment, means for searching for the non-displaying characters in the email; means for removing the searched non-displaying characters and means for tokenizing the attachment to generate a token that is representative of the attachment, wherein only the displaying characters are tokenized.

Milliken means for shows tokenizing the attachment to generate a token that is representative of the attachment ([10-13 and 51 – 53]) and

means for searching for the non-displaying characters in the email and means for removing the searched non-displaying characters, including the non-displaying comments and the non-displaying control characters ([69]), wherein only the displaying

characters are tokenized ([69]).

It would have been obvious to one of ordinary skill in the art at the time of the invention to modify the disclosure of Shipp with that of Milliken in order to better identify spam email, as at the time of Shipp's disclosure, spam email was thought “currently” not to be associated with attachments ([81]); spam and attachments are however an area for which Milliken’s more recent disclosure provides updated guidance.

Shipp in view of Milliken do not show explicitly show where the tokens are sorted in accordance with the corresponding determined spam probability value to determine a predefined number of interesting tokens, the predefined number of interesting tokens being a subset of the generated tokens.

Sahami shows selecting a predefined number of interesting tokens, the predefined number of interesting tokens being a subset of the generated tokens (pg. 4, col. 1, showing having initially “several thousand” features, then selecting 500 of said features after first sorting out features that occur fewer than 3 times (pg. 4, col. 2) and then selecting, of the remaining feature, the 500 features with the highest non-neutral probability value (pg. 6, col. 1, paragraph 1)).

It would have been obvious to one of ordinary skill in the art at the time of the invention to modify the disclosure of Shipp in view of Milliken with that of Sahami in order to more accurately identify spam email (Sahami, Abstract).

Shipp in view Milliken and Sahami thus do show selecting a subset of the generated tokens based on probability value as well as where the interesting tokens are a subset of the generated tokens (Sahami, pg. 6, col. 1, paragraph 1), but do not show

explicitly show where the tokens are sorted in accordance with the corresponding determined spam probability value.

Woitaszek shows where the tokens are sorted in accordance with the corresponding determined spam probability value (Tables 4 and 5).

It would have been obvious to one of ordinary skill in the art at the time of the invention to modify the disclosure of Shipp in view of Milliken and Sahami with that of Woitaszek in order to arrange the calculated values in a logical manner, enabling a simple method of extracting the most interesting results (as discussed by Sahami) via simply taking the top occurring results in Woitaszek's sorted list, as well as to include the abilities to integrate the spam software into a commonly used email program (Woitaszek, Abstract, pg. 1 col. 2).

36. Regarding claim 25, Shipp shows a computer-readable storage medium that includes a program, that when executed by a computer, performs the actions of

receive an email message comprising an SMTP email address ([39,43,69]), and a domain name corresponding to the SMTP email address ([39,45,46]), and an attachment, the email message further including displaying characters ([64-67]) and non-displaying characters ([57-58, 61-73]);

tokenize the SMTP email address to generate a token representative of the SMTP email address ([39,43,63])

tokenize the domain name to generate a token representative of the domain name ([22]), and determine a spam probability value from the generated tokens ([14,76]).

Shipp does not show tokenizing the attachment to generate a token that is representative of the attachment nor does Shipp show searching for the non-displaying characters in the email and removing the searched non-displaying characters, including non-displaying comments and non-displaying control characters, wherein only displaying characters are tokenized.

Milliken shows tokenizing the attachment to generate a token that is representative of the attachment ([10-13 and 51 – 53]) and

searching for the non-displaying characters in the email and removing the searched non-displaying characters, including non-displaying comments and non-displaying control characters ([69]) wherein only displaying characters are tokenized ([68-70]).

It would have been obvious to one of ordinary skill in the art at the time of the invention to modify the disclosure of Shipp with that of Milliken in order to better identify spam email, as at the time of Shipp's disclosure, spam email was thought “currently” not to be associated with attachments ([81]); spam and attachments are however an area for which Milliken’s more recent disclosure provides updated guidance.

Shipp in view of Milliken do not show explicitly show where the tokens are sorted in accordance with the corresponding determined spam probability value to determine a predefined number of interesting tokens, the predefined number of interesting tokens being a subset of the generated tokens.

Sahami shows selecting a predefined number of interesting tokens, the predefined number of interesting tokens being a subset of the generated tokens (pg. 4,

col. 1, showing having initially “several thousand” features, then selecting 500 of said features after first sorting out features that occur fewer than 3 times (pg. 4, col. 2) and then selecting, of the remaining feature, the 500 features with the highest non-neutral probability value (pg. 6, col. 1, paragraph 1)).

It would have been obvious to one of ordinary skill in the art at the time of the invention to modify the disclosure of Shipp in view of Milliken with that of Sahami in order to more accurately identify spam email (Sahami, Abstract).

Shipp in view Milliken and Sahami thus do show selecting a subset of the generated tokens based on probability value as well as where the interesting tokens are a subset of the generated tokens (Sahami, pg. 6, col. 1, paragraph 1), but do not show explicitly show where the tokens are sorted in accordance with the corresponding determined spam probability value.

Woitaszek shows where the tokens are sorted in accordance with the corresponding determined spam probability value (Tables 4 and 5).

It would have been obvious to one of ordinary skill in the art at the time of the invention to modify the disclosure of Shipp in view of Milliken and Sahami with that of Woitaszek in order to arrange the calculated values in a logical manner, enabling a simple method of extracting the most interesting results (as discussed by Sahami) via simply taking the top occurring results in Woitaszek’s sorted list, as well as to include the abilities to integrate the spam software into a commonly used email program (Woitaszek, Abstract, pg. 1 col. 2).

37. Regarding claim 26, Shipp in view of Milliken, Sahami and Woitaszek further show assigning a spam probability value to the token representative of the SMTP email address (Shipp [18,23,39,40-43], Woitaszek, Tables 4 and 5) and

assigning a spam probability value to the token representative of the domain name (Shipp [22]).

and generating a Bayesian probability values using the spam probability values assigned to the tokens (Sahami, pg.2, col. 2; pg. 4, col. 2; pg. 6, col. 1).

38. Regarding claim 27, Shipp in view of Milliken, Sahami and Woitaszek further show comparing the generated Bayesian probability value with a predefined threshold value (Sahami, pg.2, col. 2; pg. 4, col. 2; pg. 6, col. 1).

39. Regarding claim 28, Shipp in view of Milliken, Sahami and Woitaszek further show categorizing the email message as spam in response to the Bayesian probability value being greater than the predefined threshold (Sahami, pg.2, col. 2; pg. 4, col. 2; pg. 6, col. 1).

40. Regarding claim 29, Shipp in view of Milliken, Sahami and Woitaszek further show categorizing the email message as non-spam in response to the Bayesian probability value being not greater than the predefined threshold (Sahami pg. 6 col. 1).

Regarding claim 30, Shipp shows a system comprising a memory component that stores at least the following:

email receive log configured to receive an email message comprising an address (Shipp, [39,45,46]) the email message further including (Shipp, [64-67]) and non-displaying characters (Shipp, [57-58, 61-73]);

analysis logic configured to determine a spam probability value from the generated token (Shipp, [14,76]).

Shipp does not show search logic that is configured to search for the non-displaying characters in the email and remove logic that is configured to remove the searched non-displaying characters, including non-displaying comments and the non-displaying control characters, wherein only the displaying characters are tokenized, and

tokenize logic configured to generate a token that is representative of the attachment.

Milliken shows tokenize logic configured to tokenize the attachment to generate a token that is representative of the attachment ([10-13 and 51 – 53]) and

search logic that is configured to search for the non-displaying characters in the email and remove logic that is configured to remove the searched non-displaying characters, including non-displaying comments and the non-displaying control characters ([69]), wherein only the displaying characters are tokenized ([69]).

It would have been obvious to one of ordinary skill in the art at the time of the invention to modify the disclosure of Shipp with that of Milliken in order to better identify spam email, as at the time of Shipp's disclosure, spam email was thought “currently” not to be associated with attachments ([81]); spam and attachments are however an area for which Milliken’s more recent disclosure provides updated guidance.

Shipp in view of Milliken do not show explicitly show where the tokens are sorted in accordance with the corresponding determined spam probability value to determine a predefined number of interesting tokens, the predefined number of interesting tokens

being a subset of the generated tokens.

Sahami shows selecting a predefined number of interesting tokens, the predefined number of interesting tokens being a subset of the generated tokens (pg. 4, col. 1, showing having initially "several thousand" features, then selecting 500 of said features after first sorting out features that occur fewer than 3 times (pg. 4, col. 2) and then selecting, of the remaining feature, the 500 features with the highest non-neutral probability value (pg. 6, col. 1, paragraph 1)).

It would have been obvious to one of ordinary skill in the art at the time of the invention to modify the disclosure of Shipp in view of Milliken with that of Sahami in order to more accurately identify spam email (Sahami, Abstract).

Shipp in view Milliken and Sahami thus do show selecting a subset of the generated tokens based on probability value as well as where the interesting tokens are a subset of the generated tokens (Sahami, pg. 6, col. 1, paragraph 1), but do not show explicitly show where the tokens are sorted in accordance with the corresponding determined spam probability value.

Woitaszek shows where the tokens are sorted in accordance with the corresponding determined spam probability value (Tables 4 and 5).

It would have been obvious to one of ordinary skill in the art at the time of the invention to modify the disclosure of Shipp in view of Milliken and Sahami with that of Woitaszek in order to arrange the calculated values in a logical manner, enabling a simple method of extracting the most interesting results (as discussed by Sahami) via simply taking the top occurring results in Woitaszek's sorted list, as well as to include

the abilities to integrate the spam software into a commonly used email program (Woitaszek, Abstract, pg. 1 col. 2).

Regarding claim 31, Shipp shows a system comprising means for receiving an email message comprising (Shipp [18,23]) an attachment and an address (Shipp, [39,45,46]) the email message further including displaying characters (Shipp, [64-67]) and non-displaying characters (Shipp, [57-58, 61-73]) and

means for determining a spam probability values from the generated tokens (Shipp [14,76]).

Shipp does not show means for searching for the non-displaying characters in the email;

means for removing the searched non-displaying characters

means for tokenizing the attachment to generate a token representative of the attachment, wherein only the displaying characters are tokenized.

Milliken shows means for searching for the non-displaying characters in the email;

means for removing the searched non-displaying characters (Milliken, [68-70])

means for tokenizing the attachment to generate a token representative of the attachment (Milliken [10-13 and 70]); and

wherein only the displaying characters are tokenized (Milliken, [68-70])

It would have been obvious to one of ordinary skill in the art at the time of the invention to modify the disclosure of Shipp with that of Milliken in order to better identify spam email, as at the time of Shipp's disclosure, spam email was thought “currently” not

to be associated with attachments ([81]); spam and attachments are however an area for which Milliken's more recent disclosure provides updated guidance.

Shipp in view of Milliken do not show explicitly show means for sorting the tokens in accordance with the corresponding determined spam probability value to determine a predefined number of interesting tokens, the predefined number of interesting tokens being a subset of the generated tokens.

Sahami shows means for selecting a predefined number of interesting tokens, the predefined number of interesting tokens being a subset of the generated tokens (pg. 4, col. 1, showing having initially "several thousand" features, then selecting 500 of said features after first sorting out features that occur fewer than 3 times (pg. 4, col. 2) and then selecting, of the remaining feature, the 500 features with the highest non-neutral probability value (pg. 6, col. 1, paragraph 1)).

It would have been obvious to one of ordinary skill in the art at the time of the invention to modify the disclosure of Shipp in view of Milliken with that of Sahami in order to more accurately identify spam email (Sahami, Abstract).

Shipp in view Milliken and Sahami thus do show selecting a subset of the generated tokens based on probability value as well as where the interesting tokens are a subset of the generated tokens (Sahami, pg. 6, col. 1, paragraph 1), but do not show explicitly show where the tokens are sorted in accordance with the corresponding determined spam probability value.

Woitaszek shows where the tokens are sorted in accordance with the corresponding determined spam probability value (Tables 4 and 5).

It would have been obvious to one of ordinary skill in the art at the time of the invention to modify the disclosure of Shipp in view of Milliken and Sahami with that of Woitaszek in order to arrange the calculated values in a logical manner, enabling a simple method of extracting the most interesting results (as discussed by Sahami) via simply taking the top occurring results in Woitaszek's sorted list, as well as to include the abilities to integrate the spam software into a commonly used email program (Woitaszek, Abstract, pg. 1 col. 2).

41. Regarding claim 32, Shipp shows a computer-readable storage medium that includes a program, that when executed by a computer, performs at least the following:

receive an email message comprising an attachment and an address ([39,43,69]), the email message further including displaying characters ([64-67]) and non-displaying characters ([57-58, 61-73]);

and tokenizing to determining a spam probability from the generated token (Shipp, [14, 76]).

Shipp does not show performing a search for the non-displaying characters in the email;

remove the searched non-displaying characters, including non-displaying comments and non-displaying control characters, wherein only displaying characters are tokenized and

generate a token representative of the attachment.

Milliken shows performing a search for the non-displaying characters in the email;

remove the searched non-displaying characters, including the non-displaying comments and the non-displaying control characters, wherein only displaying characters are tokenized ([68-70]) and

generate a token representative of the attachment ([10-13 and 51 – 53]). ([68-70]).

It would have been obvious to one of ordinary skill in the art at the time of the invention to modify the disclosure of Shipp with that of Milliken in order to better identify spam email, as at the time of Shipp's disclosure, spam email was thought “currently” not to be associated with attachments ([81]); spam and attachments are however an area for which Milliken’s more recent disclosure provides updated guidance.

Shipp in view of Milliken do not show explicitly show where the tokens are sorted in accordance with the corresponding determined spam probability value to determine a predefined number of interesting tokens, the predefined number of interesting tokens being a subset of the generated tokens.

Sahami shows selecting a predefined number of interesting tokens, the predefined number of interesting tokens being a subset of the generated tokens (pg. 4, col. 1, showing having initially “several thousand” features, then selecting 500 of said features after first sorting out features that occur fewer than 3 times (pg. 4, col. 2) and then selecting, of the remaining feature, the 500 features with the highest non-neutral probability value (pg. 6, col. 1, paragraph 1)).

It would have been obvious to one of ordinary skill in the art at the time of the invention to modify the disclosure of Shipp in view of Milliken with that of Sahami in

order to more accurately identify spam email (Sahami, Abstract).

Shipp in view Milliken and Sahami thus do show selecting a subset of the generated tokens based on probability value as well as where the interesting tokens are a subset of the generated tokens (Sahami, pg. 6, col. 1, paragraph 1), but do not show explicitly show where the tokens are sorted in accordance with the corresponding determined spam probability value.

Woitaszek shows where the tokens are sorted in accordance with the corresponding determined spam probability value (Tables 4 and 5).

It would have been obvious to one of ordinary skill in the art at the time of the invention to modify the disclosure of Shipp in view of Milliken and Sahami with that of Woitaszek in order to arrange the calculated values in a logical manner, enabling a simple method of extracting the most interesting results (as discussed by Sahami) via simply taking the top occurring results in Woitaszek's sorted list, as well as to include the abilities to integrate the spam software into a commonly used email program (Woitaszek, Abstract, pg. 1 col. 2).

42. Regarding claim 33, Shipp in view of Milliken, Sahami and Woitaszek further show receiving an email message including a text body (Shipp [64,65]).

43. Regarding claim 34, Shipp in view of Milliken, Sahami and Woitaszek further show tokenizing the words in the text body to generate tokens representative of the words in the text body (Shipp [64,65]).

44. Regarding claim 35, Shipp in view of Milliken, Sahami and Woitaszek further show assigning a spam probability value to each of the tokens representation of the words in the text body (Woitaszek, Tables 4 and 5)

45. Regarding claim 36, Shipp in view of Milliken, Sahami and Woitaszek further show comparing the generated Bayesian probability value with a predefined threshold value (Sahami, pg. 4 col. 2).

46. Regarding claim 37, Shipp in view of Milliken, Sahami and Woitaszek further show categorizing the email message as spam in response to the Bayesian probability value being greater than the predefined threshold (Sahami, pg.2, col. 2; pg. 4, col. 2; pg. 6, col. 1).

47. Regarding claim 38, Shipp in view of Milliken, Sahami and Woitaszek further show categorizing the email message as non-spam in response to the Bayesian probability value being not greater than the predefined threshold (Sahami, pg. 6 col. 1).

Conclusion

Applicant's amendment necessitated the new ground(s) of rejection presented in this Office action. Accordingly, **THIS ACTION IS MADE FINAL**. See MPEP § 706.07(a). Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not

mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the date of this final action.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to John M. MacIlwinen whose telephone number is (571) 272-9686. The examiner can normally be reached on M-F 7:30AM - 5:00PM EST; off alternate Fridays.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Joon Hwang, can be reached on (571) 272 - 4036. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

John MacIlwinen

(571) 213 - 6095

/Joon H. Hwang/
Supervisory Patent Examiner, Art Unit 2447